Workshop Outline

Overview and Description

This workshop will focus on reproducible data practices for TMDL development. In this workshop, we will review the current landscape of data tools and resources, increase familiarity and comfort with R and Python, and create and modify code for future use.

Pre-workshop prep:

For those less familiar with R and Python, pre-meeting training resources will be circulated. Tetra Tech will be available during an office hour period to help assist those needing additional logistical support. A test script will also be provided for attendings to verify software compatibility with course materials.

Day 1 - Overview and climate data

1:00-1:30 - *Overview presentations*

- Intro: workshop goals, outline, introduce people, logistics – Jasper Hobbs, ACWA and Brenda Rashleigh, EPA
- Suitability of R vs. Python for various data tasks – Kateri Salk and Brian Pickard, Tetra Tech
- Benefits of reproducible analysis & best practices– Kateri Salk and Brian Pickard, Tetra Tech
- State perspectives – Nicholas von Stackelberg, Utah

1:30-3:00 - *Climate data acquisition harmonization* – Kateri Salk and Brian Pickard, Tetra Tech

- Data ingestion of various datasets (examples: PRISM, NOAA Atlas14)
- Data wrangling
    - o Converting native formats to data frame
    - o Transforming
    - o Summarizing data to higher levels
    - o Joining disparate data frames
- Spatial/temporal interpolation

3:00-3:10 - *10 Minute Break*

3:10-3:55 - *Examples of using climate data in modeling (45 min)*

- General Web Scraping – Eric Hettler, Wisconsin
- Scraping Climate Data from DAYMET – Eric Hettler, Wisconsin
- Developing R packages to automate data analysis – Ansel Bubel, Florida

3:55-4:00 – *5 Minute Break*

4:00-4:45 - *Data management, documentation, and export* – Kateri Salk, Tetra Tech

- Exporting data frames and plots
- Data management

4:45-5:00 - *Wrap-up and next day preview*

Day 2 - Land cover/land use

1:00-1:30 - *Overview presentations* – Brian Pickard, Tetra Tech

- Spatial modeling introduction (ESRI products + open source)
- Difference between land cover and land use

1:30-3:00 - *Acquiring and displaying spatial data (90 min)* – Brian Pickard, Tetra Tech

- Data ingestion of various datasets (examples: NLCD, NWALT, higher resolution [e.g. CBT])
- Delineating and clipping a watershed
- Summarizing watershed land use and summary information
    - Distinguishing between raster and vector data types
    - Aggregating spatial data (i.e., zonal statistics and vector summaries)

3:00-3:10 - *10 Minute Break*

3:10-3:55 - *Examples of using land cover data in modeling*

- Developing ArcGIS toolbars to automate spatial preprocessing – Ansel Bubel, Florida
- State of the science: Future land change modes (FUTURES, SLEUTH, IDRISI Products, CLUE-S) – Brian Pickard, Tetra Tech
- Open-source alternates to ESRI products – Brian Pickard, Tetra Tech

3:55-4:00 – *5 Minute Break*

4:00-4:45 - *Visualizing and understanding spatial outputs*

- Displaying spatial model outputs – Tim Wool, EPA

4:45-5:00 - *Wrap-up and next day preview*

Day 3 - Flow and water quality

1:00-1:30 - *Overview presentations* – Kateri Salk, Tetra Tech

- How empirical data are utilized in TMDL development
- Overview of water quality and flow data sources

1:30-3:00 - *Acquiring and wrangling hydrologic and water quality data*

- Flow data – Eric Hettler, Wisconsin
    - Acquisition of USGS data via DataRetrieval package
    - Basic graphing of hydrologic data
    - Creating flow-duration curves
    - Performing baseflow separation
- Water quality data – Kateri Salk, Tetra Tech
    - Acquisition of NWIS and STORET data from the Water Quality Portal (WQP)
    - Wrangling water quality data

3:00-3:10 - *10 Minute Break*

3:10-3:55 - *Examples of using hydrologic and water quality data in modeling*

- Automated post-model run data analysis & visualization with R and RESTful web services – Joey Kleiner and Rob Burgholzer, Virginia
- R Shiny apps for long term water quality in WI rivers – Eric Hettler, Wisconsin
- Utah Division of Water Quality open source data tools and interactive dashboards – Nicholas von Stackelberg, Chris Shope, and Jake Vander Laan, Utah

3:55-4:00 - 5 Minute Break

4:00-4:45 - *Visualizing and understanding model outputs*

- Graphing outputs from common water quality models (e.g., WASP) – Tim Wool, EPA Region 4
- Chlorophyll TMDL regression – Ansel Bubel, Florida

4:45-5:00 - *Wrap-up and closing comments*